HEALTH-RELATED INTERNET USE AND UPTAKE OF SEASONAL INFLUENZA VACCINE AMONG ADULTS IN THE UNITED STATES: AN ANALYSIS OF THE 2014 NATIONAL HEALTH INTERVIEW SURVEY DATA

A Thesis

by

BASHIR DABO

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Dennis Gorman
Co-Chair of Committee,	Hongwei Zhao
Committee Members,	Scott Lillibridge
	Linda L. Logan
Head of Department,	Dennis Gorman

December 2016

Major Subject: Epidemiology

ABSTRACT

Immunization against influenza is effective in preventing its seasonal outbreaks. However, lack of access to vaccines, poor knowledge and misconception of vaccine safety are among the major deterrents to the use of seasonal influenza vaccine in the US. Because of the emergence of the Internet as an important source of health information and a medium of care delivery, its use for health purposes was hypothesized to influence influenza vaccination.

In a design-based analysis of data from 36,697 participants in the 2014 National Health Interview Survey, the association between health-related Internet use and the odds of uptake of seasonal influenza vaccines among US adults was investigated. Furthermore, the difference between older adults (aged 65+) and the general adult population was evaluated in this regard. Using multiple logistic regression, the odds of influenza vaccination were examined for five variables representing use of Internet for health purposes and other covariates.

The rates of influenza vaccine use were higher in older adults (68.6%) than in the combined adult population. Among both groups, rate of performance of each of the five activities related to the use of Internet for health purposes, except use of online group chat to learn about health topics, was higher among vaccine users. Filling prescriptions online and communicating with health providers via email significantly increased the odds of accepting the vaccine for both adult groups. Searching health information on the Internet and scheduling medical appointments online were associated with significant increases in

the odds of accepting the vaccine only when all the adults were considered. Searching health information on the Internet and scheduling medical appointments online among older adults, as well as use of online group chat to learn about health topics in all the two adult populations, were not significantly associated with odds of influenza vaccine uptake.

Influenza vaccination among US adults was significantly influenced by use of Internet to obtain health information and to gain access to health services. Policy makers need to consider the importance of greater and more effective use of the Internet to successfully promote uptake of seasonal influenza vaccine.

DEDICATION

To my late father, from whom I first learned about Influenza, and my mother, who had encouraged me to take my first flu shot, ever.

ACKNOWLEDGEMENTS

The successful completion of this project would not have been possible without the relentless support of my thesis committee, which includes Drs. Dennis Gorman, Hongwei Zhao, Scott Lillibridge and Linda L. Logan. They had been the giants on whose shoulders I stood to find my way throughout this undertaking. I appreciate sincerely and acknowledge gratefully the immense contributions they made to ensure the success of this work.

My profound thanks also goes to the government of Kano state, Nigeria, under the leadership of Dr. Rabiu Musa Kwankwaso, for providing the scholarship for my study at Texas A&M University. I am grateful to Reach Care Foundation Nigeria, led by Dr. Abdulsalam Nasidi, for administering this scholarship with remarkable efficiency, and to the management of Bayero University Kano, Nigeria, for their additional support in providing me with a study fellowship.

I would also like to acknowledge the help of the US National Center for Health Statistics in making the data used in this project available for use free by the public.

Several other people, whose names may not all be mentioned here, also provided some assistance to me in this undertaking or during my study. However, I would like to register my indebtedness especially to the following:

Rabiu Sulaiman Bichi, Jean Brenda, Hugh McElroy, Samantha Payton, Brandie Taylor, Ninfa Purcell, Jackie Lavigne (from the National Institutes of Health), Erin

Schneider, Rachel P. Casey, Ashley Schmitt, Faruk Sarkinfada, Yusuf M. Sabo and Nandal Durfa.

To all others, including family and friends, who provided some support to me in whatever form during my study at Texas A&M, I say thank you.

Finally, I am grateful to God the Almighty, who is All-knowing above every possessor of knowledge, for blessing me with the wisdom and strength to carry out this project successfully.

CONTRIBUTORS AND FUNDING SOURCES

This project was supervised by a thesis committee consisting of Drs. Dennis Gorman (Chair of committee), Hongwei Zhao (Co-chair of committee) and Scott Lillibridge, all from Epidemiology and Biostatistics department, and Dr. Linda L. Logan from the School of Veterinary Medicine and Biomedical Sciences.

The data used for the entire project was produced by the National Center for Health Statistics (Atlanta, GA).

Rachel P. Casey assisted with paper editing and design of the tables presented in the Results section.

Funding for graduate study came from a scholarship provided by the government of Kano state, Nigeria (Through Reach Care Foundation, Nigeria) and a fellowship from Bayero University Kano (Nigeria).

All other effort put forth towards completion of this work was from the student, independently.

NOMENCLATURE

AAP	Asthma Action Plan
ACIP	Advisory Committee for Immunization Practices
BMI	Body Mass Index
CAPI	Computer Assisted Personal Interviewing
CDC	Centers for Disease Control and Prevention
C.I	Confidence Interval (95%)
COPD	Chronic Obstructive Pulmonary Disease
CRQ	CAPI Reference Questionnaire
DC	District of Columbia
FPL	Federal Poverty Level
GA	Georgia (State)
IRB	Institutional Review Board
MSE	Mean Square Error
NCHS	National Center for Health Statistics
NHIS	National Health Interview Survey
NH	Non-hispanic
NSR	Non Self Representing
OR	Odds Ratio
PCHMS	Personally Controlled Health Management Systems
PPS	Probability Proportional to Population Size

PSU	Primary Sampling Unit
RNA	Ribonucleic Acid
ROC	Receiver Operating Characteristic
SAS	Statistical Analysis System
SR	Self Representing
SRS	Simple Random Sampling
SUDAAN	Statistical Software for Weighting, Imputing and Analyzing Data
TSL	Taylor Series Linearization
US	United States
USA	United States of America

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1 Background

Influenza is a common and highly contagious respiratory illness which manifests as sudden onset of fever, myalgia, headache and respiratory symptoms, including cough, running nose and sore throat (Monto et. al, 2000). It is caused by three types of viruses belonging to the Orthomyxoviridae family of RNA viruses, which spread through aerosolized secretions from the respiratory tract of infected persons during coughing, sneezing or close contact (Nelson & Williams, 2014). The disease occurs in all parts of the world seasonally, and as a pandemic whenever variation occurs in the circulating viral antigen. Human Influenza virus types A and B are responsible for epidemics of the disease (Flu) in the US almost every winter (Centers for Disease Control and Prevention (CDC), 2015a). Children, pregnant women, individuals with certain medical conditions and the elderly, are among those at a particularly high risk of the disease and/or its severe complications (Barker & Mullooly, 1980).

Influenza prevention is important because of its considerable impact on the US public health system, with an estimated annual death up to about 53,000 for the three decades preceding 2007 (Thompson et al., 2010), and close to $90 billion in direct and indirect cost (Molinari et al., 2007). The Advisory Committee on Immunization Practices (ACIP) recommends vaccination of all individuals above the age of six months as the best strategy to prevent influenza (Grohskopf et al., 2015). Vaccines are available in an injectable form and as a nasal spray. One or the other could be used by all people within the recommended age range, depending on the person's age, medical conditions such as egg allergy and immunosuppression, and usage of antiviral medications (CDC, 2015b). When strains match the epidemic viral strain, the vaccine could be up to 50% to 80% effective (Fiore et al, 2010). Despite this level of protection and the support lent by cost-effectiveness analyses for mass vaccination of healthy individuals (Clements et al., 2011), utilization of influenza vaccine continues to remain low in the US. The CDC (2015c) estimated the coverage rate at 46% for all persons during the 2014-2015 influenza season. Among many predictors of influenza vaccine use, knowledge and attitude about the disease and its prevention are some of the most important (Nichol, Lofgren & Gapinski, 1992). When people have an accurate knowledge and perception with regards to influenza and the role vaccines play to lower its transmission, rate of vaccination among them tends to be higher (Martinello, Jones & Topal, 2003; Nichol, Mac Donald & Hauge, 1996).

The Internet has gained prominence over the last two decades as a significant source of knowledge on health, and as a medium of communication for delivery of health services (Covolo et al, 2013). Not only can people search factual information online from credible sources, such as the CDC, to be better informed of the safety and effectiveness of immunization against influenza, they can also use the Internet to communicate with health providers to facilitate decisions and actions with regards to uptake of the vaccine. Health-related Internet activities such as filling prescriptions or booking appointments with a health provider on the Internet may promote utilization of influenza shot/spray or other vaccines, due to their potential positive impact on access to the latter. However, information being spread by the growing online anti-vaccine movement could have the opposite effect due to its potentially negative consequences on consumer perceptions and attitudes towards vaccines. Therefore, given the diversity and high rate of use of Internet for health purposes among adults, its potential to influence influenza vaccine acceptance among the public, positively or otherwise, could be huge (Robin & Patricia, 2011). Determining the magnitude and direction of the relationship between individuals' online health-related activities and uptake of influenza vaccine, will play a pivotal role in the fight against the disease.

1.2 Problem Statement/Justification

Despite the seriousness of the need to improve uptake of influenza vaccine and the fact that Internet use for health purposes by the public may be playing an important role in determining the current rate of utilization of the vaccine in the US, research on the actual nature of the relationship between the two is limited. Studies to date have produced inconsistent results. In addition, they involved small samples and used an aggregate of online health related activity as exposure rather than specific types of health-related Internet use (Lau et al., 2012, Bourgeois et al., 2008). Therefore, a study involving a large, nationally representative sample that includes specific actions pertaining to Internet use for health purposes is necessary to evaluate the true nature of the relationship.

1.3 Study Objectives

- To study the association between use of Internet for health purposes (defined by searching health information on Internet, filling prescription online, scheduling medical appointment on Internet, communicating with health provider via email and using online group chats to learn about health topics and uptake of seasonal influenza vaccine, all within the past 12 months (of interview), among civilian, noninstitutionalized US adults, using data from the 2014 National Health Interview Survey (National Center for Health Statistics (NCHS), 2015b).

- To examine the difference(s) in the association between each of five health related Internet use activities and odds of uptake of influenza vaccine, between the older adult subpopulation (aged 65+ years) and the general adult population.

1.4 Delimitations

This research focused on the civilian, noninstitutionalized US adult population within the 50 states and the District of Columbia (DC).

1.5 Research Hypotheses

- Increased odds of taking influenza vaccine is associated with each of the following: having looked up health information on Internet; filled prescriptions online; communicated with health provider by email; and scheduled medical appointment on Internet.
- Decreased odds of using flu vaccine is associated with using online group chats to learn about health topics, all within past 12 months (of interview).

2. LITERATURE REVIEW

2.1 Use of Internet for Health Purposes among US Adults

Even two decades ago when the Internet was not as developed as it is today, its health-related application was recognized to be as diverse as medicine itself (Sunnenberg, 1997). Most of the non-professional and a significant part of the professional usage of the Internet for health purposes could be described under the term *interactive communication*, as defined by Robinson and colleagues (Robinson, Patrick & Gustafson, 1998). Online health information seeking is arguably performed with the highest frequency of all forms of consumer health-related Internet use, judging by evidence in the literature that is available on the subject. Based on a recent study, more than 62% of all US adults said they had looked up health information online within the past 12 months (Pew Research Center, 2015). According to the first nationwide household survey on Internet use as it relates to health (the first NHIS survey to include questions on the subject), many sociodemographic characteristics were found to be associated with seeking health information online (Cohen & Adams, 2011). Specifically, the highest rate of use of the Internet to search for health information, within past 12 months, was found among ages 25 to 44 years, Non-Hispanic whites, currently employed, college educated, those earning an income above 300% of the federal poverty level (FPL), and those who had private health insurance. Another study, also based on the 2009 NHIS data, reported on the effect of age (for people older than 64 years) on use of health information technology (Choi, 2011). More than a 50% reduction was found in the rate of seeking health information on the Internet, for each 10 years of increase in age above 65 years.

Considering the rapid rate of expansion of the Internet, it may be difficult to keep track of the number of websites containing information on health. Nearly a decade and half ago, an estimate of between 70,000 to 100,000 such sites was reported by researchers (Risk & Dzenowagis 2001; Cline & Haynes, 2001). Several orders of magnitude above this estimate could be in existence today considering that when a simple search on Google was performed, using the terms “human vaccine” and “human immunization”, more than 14 million results were returned for each. From these web pages, consumers of health services obtain health information in three main ways: direct searches (e.g. written material, video and audio), involvement in social/support groups, and communication with health care providers (Cline & Haynes, 2001). Direct search is probably the most common among these, but the rate at which the other two are being done could also be rising. For example, a survey showed that online communication with health providers among US adult participants rose from 7% in 2003 to 10% in 2005 (Beckjord et al., 2007). Direct access to health information on the Internet by consumers could be from recognized scientific sources (e.g. peer reviewed journals and websites of institutions like the CDC) or from unevaluated sources with questionable credibility (such as the so called “well-informed individuals”, “charlatans” and “quacks”) (Gregory-Head, 1998). The purposes of information searches are also diverse, with seeking diagnosis of a condition for self or for someone else (35% of users) and opinion about certain health interventions including vaccines (24%) among the most popular (Pew Research Center, 2015).

2.2 Association between Online Health-related Activity and Health Services Utilization

As more people continue to exploit the lower cost, flexibility, ease of access and even anonymity provided by the Internet to obtain health information and services, there could be consequences on utilization of health interventions. Authorities recognized this potential early during the emergence of Internet technology (Cline & Haynes, 2001). Also, a study in 2000 revealed that almost half of those who sought health information on the Internet for themselves said what they discovered had affected their decisions on using certain health interventions (Pew Research Center, 2000). Opinions were different among scholars and other stakeholders on how use of the Internet for health purposes could affect acceptance of health interventions by consumers (Cline & Haynes, 2001). Some believe that increased access to credible information, for example due to improved communication with health providers or visiting pages with reliable information, could increase uptake of services such as influenza vaccination. On the other hand, skeptics think that continuous proliferation of sites with damaging information, among other things, could lead to less utilization of some interventions (LaPerriere, Edwards, Romeder & Maxwell-Young, 1998). Outcomes of studies conducted to examine the actual nature of this relationship were mainly as would be anticipated, depending on the type of health related Internet activity (supportive or disruptive) used as exposure (Mortimer et al., 2015). In other words, activities that were positive and supportive of use of certain health services, were found to be associated with increased likelihood of accepting them, and vice versa. However, surprising results were obtained in some cases, suggesting that the outcome may not be as straight-forward as would be thought. For example, use of an online Personal Health Record (a positive health related interactivity) was not found to be significantly associated with use of an intervention called the Asthma Action Plan (AAP) in a randomized controlled trial involving asthmatics (Lau et al., 2015).

2.3 Health-related Internet Use and Uptake of Influenza Vaccine

Like with other health interventions, research on the association between health-related Internet use and uptake of influenza vaccine is limited. Only two peer reviewed publications, which examined the association directly, were identified in the literature research for this project. Although both studies were randomized clinical trials, their conclusions were different. One of them revealed that a type of interactive health communication called *Healthy.me* (Coiera, Lau, Anvari & Sacks, 2010) significantly increased uptake of seasonal influenza vaccine among students and staff of an Australian University (Lau et al., 2012). The other concluded that use of a web-based personal health record (a component of *Healthy.me*) did not have a statistically significant association with vaccine uptake (Bourgeois et al., 2008). With regards to health-related Internet use as the exposure, studies largely used activities with potential positive effects on uptake of influenza vaccine, such as scheduling appointments for vaccination and using online reminders to keep track of schedules. The combined effect on uptake of influenza vaccine due to activities representing Internet use for health purposes was also considered in the studies, rather than the effects of individual activities. Therefore, conclusions from the studies may only be generalized on the exposure types involved (potentially positive health-related Internet use, in a combined effect), and not by every type of Internet use for health purposes by an individual or its effect.

The inconsistency of findings from these studies, in spite of the incomprehensive nature of the exposure used, may serve as a case for conducting more research to better define the association between Internet use for health purposes and uptake of seasonal influenza vaccine. There are also other reasons that highlight an urgent need for research in this area. First, information on vaccines is commonly searched on the Internet by consumers of health care. Second, the Internet is being used by anti-vaccine campaigners to disseminate inaccurate information (Kata, 2010) or misinterpreted facts (Wolfe, Sharps & Lipsky, 2002) about vaccines to create excessive fear around their use. One simple online search fifteen years ago yielded over 300 websites with anti-vaccine content (Poland & Jacobson, 2001), suggesting that thousands of them probably exist today, based on the considerable growth in health-related Internet use since that time. Third, a majority of health information seekers online use search engines such as Bing, Google, and Yahoo rather than directly visiting those websites likely to contain credible information that could reduce, or even eliminate, the chance of being misinformed (Pew Research Center, 2015). Moreover, it has been reported that yields from such search engines for vaccines depend on the term(s) used to perform the search. For instance, the word "vaccines" was found to have yielded more websites with anti-vaccine content than when "immunization" was used (Wolfe & Sharp, 2005).

Thus, even though no scientific publication could be found that reported a negative association with uptake of influenza vaccine for any form of health-related Internet use, in the US or elsewhere, this is an important topic to address. Certain online health-related activities, such as health information searches and social media chats, could expose individuals to anti-vaccine information or opinions that are capable of influencing them to reject influenza vaccine. Therefore, studies which consider as exposure the overall activity of an individual online (not only those that are likely support vaccine use) could produce results with wider Public health implications, with regards to control of seasonal influenza.

2.4 Significance of the Study

Findings from this study could help in the determination of the need for increased efforts to improve the quality of vaccine related information on the Internet. The research could also serve as basis for evaluation of web-based intervention programs, such as Personally Controlled Health Management Systems (PCHMS), which could be used to promote uptake of vaccines. In this regard, results from this research could be compared with similar studies performed after implementation of such programs, to measure their success or failure in view of the overall usage of the Internet by the public.

3. METHODOLOGY

The study involved analysis of secondary data from the 2014 National Health Interview Survey (NHIS), a nationally representative cross-sectional survey (NCHS, 2015a).

3.1 The NHIS

The National Health Interview Survey is one of the major surveys executed by the National Center for Health Statistics (NCHS), a division of the US CDC established by the National Health Survey Act (1956) to collect data on a wide range of topics on the health of civilian, noninstitutionalized US population, and has been a major source of knowledge on the health of the population since its commencement in July 1957. It is conducted annually and, as its major advantage, has the ability to classify the health information it collects by several socioeconomic and demographic characteristics within the US population. Contents of the survey are being revised periodically since its inception, with the current format being in place since the last major revision in 1997 (NCHS, 2015b).

3.1.1 Survey Design

NHIS is an interview based cross-sectional household survey, aimed at the non-institutional part of the US population residing within the country at the time of interview. Its structure is complex, with a multistage area probability design that enabled households and group quarters (e.g. school dormitories) to be sampled representatively. Based on the sampling format used for the 2014 survey, a sample of 428 primary sampling units (PSUs) was selected from around 1,900 PSUs (nationally) that were geographically defined. PSUs (clusters) were described as a county, a small collection of counties that are contiguous, or a metropolitan statistical area. Their selection was stratified geographically based on certain criteria, including region and state. Within nearly each stratum, at least two PSUs were selected randomly without replacement. Selection probability was proportional to the size of the population (PPS). These types of PSUs are known as non-self-representing (NSR) or non-certainty PSUs, and consisted of relatively small populations. However, certain PSUs included very large populations (e.g. Los Angeles metropolitan area) and, for operational efficiency and cost-effectiveness, they were selected with certainty. These are referred to as self-representing (SR) or certainty PSUs. In the end, each of the 50 states and the District of Columbia were represented by some PSUs. Area segments and permit segments were used within a PSU as second stage units. The former was defined geographically. However, permit segments were described temporally as “housing units built following the 2000 census.” Two non-superimposing parts make up the NHIS sampling frame. These include the area frame (the entire area segments) and permit frame (sum of the permit segments). The US Census Bureau provides NCHS with address listings from which the NHIS households are sampled at random. This list is independent of census operations. However, commercial listings have been included for a small portion of the 2014 survey.

To gain higher precision of estimates for blacks, Hispanics and Asians during analysis of the survey data, individuals from these groups who were 65 years or older, were deliberately oversampled. Details of this oversampling process are described on the

NHIS website. In addition, through a process called PSU augmentation (Parson et al., 2014), the 2014 (and 2013) surveys consisted of more clusters than their predecessors.

3.1.2 Study Population and Inclusion Criteria

The survey, generally, was aimed at US civilian adult subpopulation that is noninstitutionalized. For the purpose of this analysis, the sample includes adults (aged 18 years and above) living in households and group quarters (e.g. school dormitories) within the 50 states and the DC. Excluded were active-duty members of Armed Forces, people in correctional facilities and those in health institutions. However, Armed Forces personnel who were actively serving were included, when at least one member of the household to which they belong was a civilian. They were eligible to participate when randomly selected but were given a weight of zero in the final datasets. Therefore, they would not be included in the estimation of standard errors and other population parameters during analysis. There were 347 such participants in the 2014 survey.

3.1.3 Sample Size and Response Rate

Determination of optimal sample size for a complex survey like the NHIS involves a delicate balance between the desired accuracy of estimates from the data collected (reflected in standard error rate, 95% confidence interval, unit attrition rate and design effect), feasibility as well as the cost of executing the survey (United Nations, 2008; Groves, 2004).

NHIS has not provided the details of values for the determining parameters and the cost implications considered in reaching the initial target of households for the survey.

Note that the 2014 survey sampled 28% more households compared to the typical NHIS target of close to 45, 000 before PSU augmentation was introduced in 2011.

Due to the design nature and the way complex surveys are conducted, it is very difficult (if not impossible) to have complete responses. Attrition could occur at each stage of the survey exercise. For example, a selected household could be found to have no occupants present throughout the survey period, ineligible participants (e.g. active duty Armed Forces personnel) or, ultimately, occupants not willing to participate in the survey despite eligibility. The same problems could be encountered at the family level and when individual level respondents (adults, in this case) are selected from eligible families. The "adult public use data file" contains records for 36,697 respondents from 45,597 families and 44,552 households. The conditional response rate (number of completed adult interviews divided by the sum of eligible adults) was 80.5%. However, when this is multiplied by the family level response rate (73.1%) to reflect nonresponse thereof, the new rate, called the "unconditional response rate," becomes 58.9%. Note that post-survey adjustments of sample weights could have eliminated bias due to this non-response, or at least minimized it (Levy and Lemeshow, 2013). Assurance has been given by the NHIS that estimates at the national and regional levels are reliable using the final samples in the public release data files (Parsons et al., 2014)

3.1.4 Participants Selection

Sample adults were selected randomly, one from each family in the sampled household. Households were defined as occupied housing units, while a family refers to a person or group of related individuals who live in one household. In some instances, individuals that were not related but were sharing the same household (such as unmarried couples) were regarded as family.

3.1.5 Data Collection

Data was collected using a questionnaire based (face-to-face) home interview. The data collection process was the responsibility of the US Census Bureau under a contract with the NHIS. Approximately 750 well-trained interviewers (field representatives) conducted these interviews across the nation (Parsons et al., 2014).

3.1.5.1 The 2014 NHIS Questionnaire: Study Exposure and Outcome

The bulk of the questionnaire for the 2014 survey had been in use since 1997 and consisted of a core (which is constant for every survey year) and a supplement. The core covered basic information on the status of participants' health, use of health services, and behaviors related to health. The supplement consisted of a range of health conditions and related events including this study's outcome variable (uptake of influenza shot or nasal spray in the past 12 months) and five main exposure variables (looked up health information on the Internet; filled a prescription on the Internet; scheduled a medical appointment on Internet; communicated with health provider by email; and used online group chats to learn about health topics, all within past 12 months) for this study. All were coded as 1 and 0, indicating performance or otherwise of the activity being measured by the variable, respectively. Note that two variables are present in the 2014 adult dataset, for uptake of seasonal influenza vaccine, i.e. uptake of influenza intramuscular injection (more than 99% of vaccine users) and use of influenza nasal spray inoculation. They were merged into one variable for use in the analysis as the outcome variable. As mentioned earlier, both the core and supplement questionnaires were completed using household interviews. Overall, more than 700 variables have been included in the sample "adult public data file."

3.1.5.2 Interview Procedure

The interview was conducted using the CAPI (Computer Assisted Personal Interviewing) software. CAPI reference questionnaires (CRQ) were made available on the computer screens of interviewers, who enter subjects' responses directly into a computer. However, where necessary, per a respondent's request or when conditions did not allow direct home visit before expiration of survey period, interviews or follow ups (as the case may be) were performed via telephone. Moreover, in 488 cases, a proxy responded on behalf of the sampled adult. Once a household was sampled for the survey, an advance letter, which provided details about the survey, was sent to the occupants before interviewers visited the household.

3.1.5.3 Data Quality Control

To ensure smooth conduct of the survey and collection of high quality information, the NHIS interviewers were supervised by career civil servants at the Census Bureau who were selected competitively through testing and received refresher trainings while the survey was being conducted. Also, performance of the process was monitored monthly to observe non-response and telephone interview rates and to ensure that these rates do not pass specified limits. Additionally, the CAPI contained features to ensure consistency of responses. For example, the software could direct the interview based on previous responses, check the consistency of an entry by comparing it to other responses for the individual, and save the information entered directly in to a survey file.

3.1.6 Post Survey Modifications

3.1.6.1 Sampling Weight Modification

Based on the probability proportionate to population size (PPS) sampling, respondents were assigned weights to reflect the number of people they represented in their respective PSUs. This initial weight, called the base weight, was a product of inverses of probabilities of selection at each stage of the survey, for each individual (Moriarty & Dahlhamer, 2012). Some procedural and technical considerations made it necessary to adjust these initial weights for the participants (Parson et al., 2014). The modifications were as follows:

1. Household Non-response Adjustment

Involved multiplying the base weight by a factor equal to the ratio of all eligible households to that of surveyed households in a PSU, up to a maximum of 2.0. Where oversampling was done, this factor was multiplied by the oversampling factor for the particular PSU.

2. Ratio and Adult Non-response Adjustments:

These were applied to person level and individual level weights, respectively. Ratio adjustments (first and second stages) were meant to correct the potential underestimation of mean square error (MSE) estimates due to the use of ratio estimators (compared to base-weight estimators). Ratio adjustments were also performed to reduce bias due to systematic under-coverage during the survey, e.g. where populations were difficult to sample or hard to reach. Adult non-response adjustment was done in between the first and second ratio adjustments.

a. First Stage Adjustment:

This was aimed at minimizing the between PSU variance of sampling variation for the NSR PSUs. Factors used for the "Sample Adults" adjustment were created by the Census Bureau for each of the eight "race and residence and ethnicity" classes across the four census regions of the US.

b. Adult Non-response Adjustment:

It was performed to reduce bias due to non-response for the sub-sampled adults in the families. The method for this adjustment was similar to the one used for the household non-response adjustment. The computation was done following the inflation factor application for choosing the sample adult.

c. Second Stage Ratio Adjustment (Poststratification):

This involved the application of a second ratio factor called the poststratification adjustment weight. Its purpose was to ensure agreement between estimates from NHIS data and those for controls selected from the general population. The Census Bureau prepared the sample used as the control, which was compared with 100 age-sex-race and ethnicity classes of the civilian, noninstitutionalized US population.

3.1.6.2 Post Survey Data Editing

Despite the various data quality assurance steps taken during survey data collection, other measures were necessary after the survey was completed. For example, validity of responses to questions was checked and spurious entries were either followed up for confirmation or deleted. Also, in the case of blank responses, those that could be filled using other information provided by the respondent were completed, while the rest were left as they were.

3.1.6.3 Data Recodes and Suppression of Variables

To fulfill the stringent confidentiality requirements before NHIS data is released publicly, personal identifying variables had to be either recoded or suppressed (not released for use openly by the public, grossly altered or merged with other variables). The type of treatment for each variable was a delicate balance between the sensitivity of the information it contained, its characteristics and how important it is to statistical analysis. The variables that underwent this sort of modification among those used in this study include age (top coded at 85) and strata (certain SR PSUs were masked by merging them with other PSUs). Body mass index and family income-to-poverty ratio variables were indirectly affected. The former was derived from the "weight" and "height" variables while the latter from the "family income" variable, all of which were either top coded or masked by use of intervals instead of points. Personal income and detailed occupation variables, which could be important confounders in the association this study examined, were completely removed from the public files.

3.1.7 Ethical Considerations

3.1.7.1 Confidentiality

NHIS data were collected under agreement that personally identifiable data is to be used only for statistical and research purposes. This is detailed in the Public Health Service Act (section 308d) and the Confidential Information Protection and Statistical

Efficiency Act (section 512b). Violation of the latter act constitutes a class E felony and is punishable by a prison term up to five years or a fine up to $ 250,000, or both. Assurance of adherence to the provisions of the acts (including all steps to be taken) was communicated to sampled households in the “advance letter” mentioned previously. The letter was also given to the participants just before the interview. The data recording and variable suppression mentioned above are examples of steps taken to ensure confidentiality of information. Moreover, NCHS has provided conditions upon which data users must agree before downloading the data from the NHIS website for analysis.

3.1.7.2 Consent

Before the NHIS interview commenced, participants were notified about the voluntary nature of the survey including their right to stop participation at any time. They were given an informed consent to sign prior to interviewing.

3.1.7.3 Compensation

No compensation was given to respondents for participating in the survey

3.1.7.4 Institutional Review Board (IRB) Clearance

According to Texas A&M University Division of Research’s guidelines, use of the NHIS data for study, as it is presented for public use on the NHIS website, does not require further review by the University’s IRB (Texas A&M University Division of Research, 2012).

3.2 Data Analysis

The public use sample adult file (2014) on the data producer’s (NHIS) website was stored in ACI format. It was downloaded and adapted for use in Stata (version 14), the software used for this analysis. Stata commands for the conversion were also available on the same website. Choice of the software for this work was informed by the great acclamation it received for versatility and ease of use in the analysis of survey data (Levy & Lemeshow, 2013). Since the data were not collected by simple random sampling (SRS) and the independence of observations could not be guaranteed due to clustering, design based analysis was performed. This method of analysis is based on the survey design features (final weight, strata, and cluster) and does not assume any form of known distribution or randomness for the data (Lumley, 2004). It ensured standard errors were not underestimated and the chance of type 1 error was minimized, compared to when the data is treated as SRS. To describe the data, means were obtained for age and body mass index (BMI) while proportions were computed for the binary and categorical variables. A multiple logistic regression model was built for hypothesis testing, using the outcome variable, main exposure variables and other selected covariates. Odds ratios (OR) were used as the measure of association. The 2014 NHIS data has been adapted for variance estimation using either linearization (Taylor Series Linearization) or replication methods (Balanced Repeated Replication, Jackknife and Bootstrapping) (Parson et al., 2014). Taylor Series Linearization (TSL) was used for this report.

3.2.1 Data Examination and Cleaning

Before starting the model building, certain measures were taken to ensure appropriateness of the downloaded dataset for the analysis. First, the authenticity of the file and its completeness were tested based on a procedure recommended by the NHIS (NCHS, 2015b). Second, the codes for the cluster, strata and weight variables were identified to make sure they were consistent with the data producer's documentation. In addition, the distribution and scaling of the weight variable were examined for outliers, missing values etc., which could affect the stability and accuracy of estimates. Third, a final data file was created that contained only the variables needed for the analysis. Note that this followed after the variable selection step described below. Finally, rates and patterns of missing data (item non-response) were explored.

3.2.2 Initial Selection of Candidate Covariates for the Logistic Model

The NHIS is a multipurpose survey on the health of its target population and the sample adult data file included over 700 variables. An inspection of the variables layout suggested that only few of them could have some real life relationship with influenza vaccine uptake and the five health related Internet use variables, to be considered for inclusion in the logistic regression model as potential confounders. Therefore, in selecting these covariates, the following procedure was adopted. First, all five health-related Internet use variables were automatically chosen because they form the main exposures for this study. Second, existing literature on confounding of the association between exposures and the outcome was used, with the help of directed acyclic graphs (Weng, Hsueh, Messam & Hertz-Picciotto, 2009) to select other covariates. Moreover, because it was not possible to examine all the literature available on the subject within the study's time limit, experts were consulted in deciding if any among the remaining covariates lacking available literature could be reasonably considered as potential confounders.

Overall, 28 covariates were identified as potential confounders. They include age (0=18-29, 1=30-49, 2=50-64, 3=65-74, 4=75-84, 5=85+), gender (0=Male, 1=Female), race/ethnicity (0=White Non-Hispanic, 1=Black/African American, 2=Hispanic, 3=Asian, 4=Other), region (0=Northeast, 1=Midwest, 2=South, 3=Midwest), marital status (0=Married/living with partner, 1=previously married, 2=never married/not living with partner), proxy as respondent (0=no, 1=yes), family income-to-federal poverty level ratio (0=2 or above, 1= 1 to 1.99, 2=less than 1), work status past 12 months (0=Had job, 1=Had no job), healthcare job/volunteer preceding 12 months (0=No, 1=Yes), trouble finding health provider past 12 months (0=no, 1=yes), time since last contact with health professional (0=1 year or less, 1= more than 1 year but less than 2 years, 3= more than 2 years but less than 5 years, 4= 5 years or never), level of satisfaction with health care past 12 months (0=very satisfied, 1=somewhat satisfied, 2=not satisfied), smoking status (0=never, 1=current, 2= former), alcohol consumption (0=never, 1=former, 3=current), BMI (continuous in kg/m^2), ever been told you have hypertension (0=no, 1=yes), ever told you had/have the following: high cholesterol (0=no, 1=yes), heart disease/condition (0=no, 1=yes), stroke (0=no, 1=yes), chronic obstructive pulmonary disease (COPD) (0=no, 1=yes), asthma (0=no, 1=yes), ulcer (0=no, 1=yes), cancer (0=no, 1=yes), diabetes (0=no, 1=yes), kidney failure (0=no, 1=yes), arthritis (0=no, 1=yes), chronic liver disease (0=no, 1=yes), and if ever had pneumonia shot (0=no, 1=yes). Responses coded as “0” were used as the reference for all analyses. Note that to make them more suitable for model building, some variables were re-categorized or transformed, as such they are not presented here as they are in the original data set. All of these except age and family income-to-poverty ratio were categorized to ensure balance of observations within cells. See section 3.2.3.3 for details on transformation for these latter two covariates.

3.2.3 Model Construction

Before a final logistic model is defined for inferences, many tentative models would be specified; their parameters (and standard errors) estimated and evaluated; and their fitness checked (Heeringa, West & Berglund, 2010). For this purpose, Heeringer et al.'s recommendation, based on Hosmer and Lemeshow's incremental method for specifying the initial logistic model (Hosmer, Lemeshow & Cook, 2000), was adopted in the following steps:

3.2.3.1 Univariate Analysis

Univariate logistic regression was performed with the outcome variable vs. each of the main exposure variables and other covariates identified in section 3.2.2. Rao-Scott's F-statistics (Rao & Scott, 1987) was used to test the significance of the univariate association. All the covariates selected previously were included as candidates for the initial multivariate (main effect) model, based on p-value (for the F-test) less than 0.25.

3.2.3.2 Multivariate Analysis

In this step, the outcome variable was regressed over all the covariates (selected in section 3.2.3.1 above) simultaneously, in a multivariate main effect logistic model. The significance of each covariate in the multivariate model was then tested, one at a time, using adjusted Wald's F-test (Heeringa, West & Berglund, 2010). Those found significant based on Wald's test p-value (at 95% significance level) were retained in the model, while the insignificant were taken out. Starting from the least significant covariate, this process was repeated until no more covariate was left that was insignificant, unless it was one of the five variables for health related Internet use. The resulting multivariate model was considered the final ‘main effect’ model before testing interactions.

Note that the adjusted Wald’s F-statistics tests the null hypothesis that the parameter coefficient for the covariate (being tested for significance in the model), is not significantly different from zero, given other covariates are in the model. The opposite forms the alternate hypothesis. Therefore, failure to reject the null hypothesis (p-value > 0.05) means the covariate could be dropped from the model.

3.2.3.3 Checking Linearity Assumption for the Continuous Covariates

Age and family income-to-poverty ratio were the only continuous covariates that were significant in the main effect model. The former satisfied the linearity assumption required for logistic regression when checked using scatter plot (with Lowess line) of the outcome variable versus age. Nevertheless, to facilitate ease of interpretation of results and make comparisons with other studies more meaningful, age was categorized according to common practices in some of the key sources of literature reviewed on US health-related Internet use and influenza vaccine uptake (CDC, 2015c; Choi 2015; Lenhart, Purcell & Smith, 2010). Interestingly, all logistic models built using age as a categorical covariate were very similar to the ones in which the variable was linear. The family income-to-need ratio variable, however, had no linear relationship with the outcome variable even after several different transformations were considered. Therefore, it was also categorized. In doing so, the FPL was considered as the first division point so that difference in vaccine acceptance between those with family income below it and those above it, up to 200%, would be observed. No further classification beyond the latter point was done based on the assumption that since influenza vaccine is relatively affordable, there may not be any significant difference in its utilization based on income alone, when the family income was above 200% of need.

3.2.3.4 Checking for Interactions between the Covariates

Significance to the main effects model of all first order interactions was tested using adjusted Wald's test, just as the covariates in the initial multivariate model were tested (Heeringa, West & Berglund, 2010). Significant and scientifically plausible interactions were added to the main effects (reduced) model to make the final (fitted) model, based on which inferences were drawn.

3.2.3.5 Missing Data Analysis

About 94% of all observations were included in the final models for both the general adults and the older adults, when listwise deletion was performed. This rate of item non-response was unlikely to have affected the estimates significantly (Heeringa, West & Berglund, 2010). As a result, attempt to improve their precision using multiple imputation (He, Zaslavsky, Landrum, Harrington & Catalano, 2009; Schenker et al. 2006) was not made. Moreover, fulfilling all of the important requirements for more accurate estimates in a multiple imputation model for a large dataset like the NHIS, would have required more time than was available for this study (Sterne et al, 2009). However, it is important to note that the original family income variable (not released publicly by the NHIS) had more than 20% missing values which were multiply imputed by the data producer. As a result, the family income-to-poverty ratio variable derived from it was released in a separate data file containing five datasets. The decision to use the family income-to-need ratio variable as categorical in the logistic models negated the need to recognize it as imputed (combining all five data sets) during odds ratio estimations (NCHS, 2015a), since all the different imputed forms yielded exactly the same distribution across the categories used. Hence, one completed data set was selected and categorized for use in the analysis.

3.2.3.6 Model Evaluation

The complex design of surveys makes it inappropriate to apply model diagnostics used for data collected by SRS (e.g. Deviance chi-squared tests, Hosmer-Lemeshow test of goodness-of-fit and area under Receiver Operating Characteristic (ROC) curve) to evaluate the goodness of fit of regression models built using survey data. Adaptation of these techniques for use in common software, such as Stata, SAS and SUDAAN, to evaluate survey logistic regression models are still under development. Nevertheless, an extension of the standard Hosmer-Lemeshow goodness-of-fit test (Archer, Lemeshow and Hosmer, 2007) has been in use in Stata (*svylogitof*) for some time. It was employed in testing how well the logistic model fitted the data.

3.2.4 Subpopulation Analysis

A separate model was built for older adults (aged 65 years or older) following the same steps used for the primary analysis performed for all adults. This was in view of significant difference observed from descriptive analysis of the data used in this study, in both health related Internet use and uptake of influenza vaccine between older adults and their younger counterparts. Other sources also reported older adults to have used more influenza vaccine (CDC, 2015) and less Internet for health purposes (Choi, 2011) than adults who were younger.

3.2.5 Model Interpretation

Statistical significance of the association between exposure and other covariates and the outcome was inferred using p-values and 95% confidence intervals for the odds ratios. Details of this are presented in the Results section.

4 RESULTS

4.1 Influenza Vaccine Uptake by Health-related Internet Use, Sociodemographic and Health Characteristics

The 2014 NHIS sample adult file contains data for 36,697 individuals, representing a population estimate of 239,688,457 civilian, non-institutionalized US adults. Of these, 8,644 (estimate = 44,934,480) were older adults (aged at least 65 years). An estimated 41.4% of all the adults and 68.6% of the older adults reported taking influenza vaccine (shot or spray) within the past 12 months (Table 1). Most of influenza vaccine utilization was in its injectable form, with the nasal spray uptake (not shown on the table) reported by less than 1% of vaccine users. Within the same period, 42.62% of all the adults and 28.6% of older adults said they had searched heath information on the Internet. However, less than 10% of the population in each case, was estimated to have performed each of the four other main exposure activities, namely: filling prescriptions on Internet (6.7% of all the adults, 6.8% of older adults); scheduling medical appointments on the Internet (6.6% of all the adults, 3.2% of older adults); communicating with healthcare providers by email (7.7% of all the adults, 6.0% of older adults); and using online group chat to learn about health topics (2.5% of all adults, 1.5% of older adults). Weighted mean age was 47 years (C.I: 46.6, 47.3) for all the adults and 73.8 years (C.I: 73.6, 74.0) for the older adults. Majority in each case were Non-hispanic white, from the southern US and within the 30-49 and 65-74 age categories, respectively. Also, 51.8% of all the adults and 55.6% of the older adults were females. Item non-response rate for all the variables included in the analysis ranged from 0% to 4.39%, based on the unweighted estimates. Significant differences were observed between those who had taken influenza vaccine and those who had not, with regards to all forms of health-related Internet use (except use of online group chat to learn about health topics), as well as certain sociodemographic and health characteristics. Among the general adults group and compared to non-users, influenza vaccine users had higher proportion of performance of each of the four activities that constitute health-related Internet use for which significant differences were observed according to vaccine use. Influenza vaccine users also had a higher proportion of females, previously married individuals (widowed/divorced/separated), Non-hispanic white and Asian races, individuals from the Northeast and Midwest, those who had worked or volunteered in a health setting during the 12 months preceding interview, former smokers and former alcohol consumers, those who had contact with health provider within less than 1 year, those who were very satisfied with health care within previous 12 months as well as those who had ever taken pneumonia shot. Surprisingly, a higher proportion of adults who had not used influenza vaccine had a job the previous 12 months versus those who had taken the vaccine. Rates of all the chronic diseases were significantly higher for vaccine users when the combined adult population was considered. .

Table 1. Estimated (Weighted) Adult Population Percentages of Health-related Internet Use and Other Characteristics, According to Influenza Vaccine Uptake, USA, 2014

	Distribution estimates	Influenza vaccine (shot or nasal spray) uptake past 12 months					
		General adults (N=239.69 million)			Older adults (65+) (N=42.2 million)		
Exposure	Overall Percentages General adults(Older adults)	Yes (41.4%)	No (56.9%)	P*	Yes (68.6%)	No (29.4%)	P*
Looked up health information on Internet	42.6(28.6)	46.3	41.1	0.000	31.3	24.3	0.000
Filled prescription on Internet	6.7(6.8)	10.2	4.4	0.000	8.3	3.7	0.000
Scheduled medical appointment on Internet	6.6(3.2)	8.7	5.3	0.000	3.8	2.0	0.003
Communicated with health provider via email	7.7(6.0)	10.8	5.6	0.000	7.3	3.3	0.000
Used online group chats to learn about health topics	2.5(1.5)	2.6	2.4	0.310	1.6	1.5	0.920
Age (years)				0.000			0.000
18-29	21.3	13.5	27.0		……	……	
30-49	34.0	27.5	38.9		……	…...	
50-64	25.9	27.9	24.5		……	……	
65-74	11.0(58.7)	17.5	6.3		56.4	64.7	
75-84	5.6(29.6)	9.6	2.6		30.8	26.8	
85+	2.2(11.7)	4.0	0.8		12.8	8.6	
Sex				0.000			0.870
Male	48.2(44.4)	43.3	51.7		44.5	44.4	
Female	51.8(55.6)	56.7	48.3		55.5	55.8	
Marriage				0.000			0.030
Married/cohabiting	75.4(60.8)	74.7	76.0		62.0	57.8	
Previously married	17.2(37.2)	20.3	14.8		36.2	39.9	
Never married/not cohabiting	7.3(1.8)	4.8	9.1		1.63	2.2	

Table 1. Continued

	Distribution estimates	Influenza vaccine (shot or nasal spray) uptake past 12 months					
Exposure	Overall percentages General adults(Older adults)	General adults (N=239.69 million)			Older adults (65+) (N=42.2 million)		
		Yes (41.4%)	No (56.9%)	P*	Yes (68.6%)	No (29.4%)	P*
Race/Ethnicity				0.000			0.000
NH White	65.7(78.4)	71.6	61.2		80.9	72.3	
NH Black	11.6(8.5)	9.5	13.2		7.0	12.1	
Hispanic	15.3(7.7)	11.2	18.2		6.7	10.1	
Asian	3.9(4.1)	6.0	5.2		4.3	3.8	
Other	1.4(1.3)	1.6	2.1		1.1	1.7	
Region				0.001			0.920
Northeast	17.3(19.6)	18.4	16.9		20.1	19.0	
Midwest	23.0(22.5)	23.7	22.3		22.3	22.8	
South	37.2(37.3)	36.7	37.9		37.2	37.6	
West	22.5(20.6)	21.2	23.5		20.5	20.6	
Had job past 12 months	67.5(21.8)	59.3	73.4	0.000	20.1	26.1	0.000
Worked/volunteered in health setting (past 12 months)	8.1(3.6)	12.5	5.1	0.000	3.8	3.5	0.590
Alcohol drinking				0.000			0.008
Never	20.9(24.7)	20.5	21.3		24.0	26.9	
Former	13.4(23.9)	15.8	11.8		23.4	25.8	
Current	64.2(49.7)	63.1	65.8		51.9	46.3	
Smoking				0.000			0.000
Never	61.0(52.2)	60.8	61.9		51.8	54.4	
Current	16.7(8.5)	11.5	20.5		7.2	11.3	
Former	21.8(38.6)	27.6	17.8		40.9	34.2	
Family income-to-FPL ratio				0.000			0.000
≥ 2	67.3(68.4)	73.1	63.0		71.1	61.6	
1 to 1.99	18.8(22.2)	16.4	20.7		20.7	26.2	
< 1	13.9(9.4)	10.5	16.2		8.2	12.2	
Response by proxy	1.6(2.8)	1.8	1.4	0.000	2.5	3.4	0.110

Table 1. Continued

Exposure	Distribution estimates: Overall percentages General adults(Older adults)	Influenza vaccine (shot or nasal spray) uptake past 12 months: General adults (N=239.69 million) Yes (41.4%)	No (56.9%)	P*	Older adults (65+) (N=42.2 million) Yes (68.6%)	No (29.4%)	P*
Level of Satisfaction with health care				0.000			0.000
Had no care	11.7(2.4)	3.4	18.0		0.7	6.6	
Very	59.0(73.4)	70.5	52.2		77.8	68.1	
Somewhat	22.2(18.6)	21.3	23.4		18.1	20.6	
Unsatisfied	4.8(3.2)	4.0	5.5		2.7	4.2	
Had trouble finding health provider past 12 months	2.4(1.8)	2.2	3.0	0.000	1.7	1.9	0.640
Time since last talked to a health professional				0.000			0.000
Less than 1 year	82.0(93.9)	93.6	75.4		97.6	88.6	
1-2 years	7.3(2.0)	3.7	10.1		1.3	3.8	
2-5 years	5.0(1.3)	1.4	7.8		0.5	3.2	
>5years/Never	4.2(1.6)	1.3	6.4		0.6	4.1	
Had pneumonia shot	74.5(58.7)	36.8	9.7	0.000	72.5	30.3	0.000
Arthritis	22.5(48.7)	31.0	16.2	0.000	51.4	42.4	0.000
Asthma	12.8(10.9)	13.9	12.1	0.000	11.8	9.1	0.005
BMI (kg/m2)				0.020			0.980
Underweight	1.8(1.9)	1.7	1.91		1.9	1.9	
Normal	33.0(31.0)	32.2	34.1		31.3	31.2	
Overweight	33.3(36.6)	33.5	33.6		36.9	37.0	
Obese	28.3(26.2)	29.6	27.7		26.7	26.0	
Cancer	8.5(23.1)	12.6	5.5	0.000	24.6	19.4	0.000
Chronic liver disease	1.1(1.6)	1.5	0.8	0.000	1.6	1.6	0.960
COPD	3.1(8.1)	4.8	1.9	0.000	9.0	6.0	0.001
Diabetes	10.7(20.5)	15.4	7.3	0.000	24.9	19.6	0.000
Heart disease	7.9(18.4)	10.8	5.7	0.000	20.1	14.1	0.000
High blood pressure	31.0(62.4)	40.7	23.8	0.000	65.8	54.3	0.000

Table 1. Continued

Exposure	Distribution estimates: Overall percentages General adults(Older adults)	Influenza vaccine (shot or nasal spray) uptake past 12 months: General adults (N=239.69 million) Yes (41.4%)	No (56.9%)	P*	Older adults (65+) (N=42.2 million) Yes (68.6%)	No (29.4%)	P*
High cholesterol	28.5(54.7)	38.4	21.4	0.000	57.5	47.9	0.000
Kidney failure	1.9(5.1)	2.7	1.2	0.000	5.3	4.5	0.280
Stroke	2.6(7.2)	3.9	1.7	0.000	7.8	5.9	0.023
Ulcer	6.7(11.3)	8.2	5.6	0.000	11.8	10.2	0.090

* *P* indicates probability for difference between vaccine users and nonusers, from design-based F-test

Among adults aged 65 years or older (Table 1), those who had accepted influenza vaccine within past 12 months also had a significantly higher proportion of individuals who had performed each of all five activities representing Internet use for health purposes, except use of online group chat to learn about health topics. Compared to non-users, older adult vaccine users were more likely to be at least 75 years old, married/cohabiting, Non-hispanic white or Asian, without job past 12 months, current alcohol consumers, former-smokers, from a family earning more than twice the federal poverty income threshold as well as very satisfied with health care in the 12 months prior to interview. Older adults who had accepted influenza vaccine within past 12 months were also more likely to have had contact with a health professional within less than a year of interview, to have taken a pneumonia shot at least once in their life time as well as to be diagnosed with any of the chronic conditions used in the analysis, except chronic liver disease, kidney failure and ulcer, which were found to have no significant difference in distribution between vaccine users and non-users.

Further analysis (not shown in Table 1) revealed that among all adults, those who performed the most common health-related Internet activity (looking up health information on the Internet) had lower average age and were more likely to be female, white, married or living with a partner, employed as well as working in health care setting. They were also more likely to, within the past 12 months, have had trouble finding a health provider, be more satisfied with health care and have contacted a health professional. Performing health information search on the Internet was also related to higher proportion of adults who never had pneumonia shot, those with family income at least twice the federal poverty level and those with asthma or cancer. On the other hand, higher proportion of adults with diabetes, hypertension, stroke and COPD was found among those who had not looked up health information on the Internet, compared to those who had performed the activity. Individuals who had used online group chat to learn about health topics also had higher proportion of those with almost every characteristic mentioned above for those who searched health information online. The few exceptions were that online group chat users were more likely to be dissatisfied with health care within past 12 months, and that they had higher proportion of those who had ever taken pneumonia shot, compared to non-users. The differences with respect to ever taking pneumonia shot and marital status, between users and non-users of online group health chats, were not statistically significant.

4.2 Association between Health-related Internet Use and Uptake of Influenza Vaccine

A total of 34,370 individuals were included in the final data analysis. A multivariate logistic regression model was fit for all the adults (Table 2), adjusted for age, gender, race/ethnicity, region, marital status, family income-to-poverty ratio, smoking status, work status, work/volunteer in a health setting, level of satisfaction with health care, time since last seen/talked to a health professional and statuses of smoking, high cholesterol, hypertension, arthritis and chronic liver disease. These complete cases represent about 94% of the total population estimate for adults. The results for the model (F:50, 251 = 66.16, $P<0.0001$) show that increased odds of influenza vaccine uptake within 12 months prior to interview was associated with looking up health information on the Internet (OR=1.11, C.I: 1.02, 1.20), filling prescriptions on Internet (OR=1.48, C.I: 1.25, 1.75), scheduling medical appointments on Internet (OR=1.18, C.I: 1.02, 1.38), and communicating with health providers by email (OR=1.27, C.I: 1.09, 1.48). Use of online group chat to learn about health topics, however, was not significantly associated with odds of using the vaccine (OR=0.86, C.I: 0.70, 1.07). All the sociodemographic and other covariates not related to Internet use included in the analysis were significant to the logistic model. Increase in odds of accepting influenza vaccine was associated with being female; in older age category (vs 18-29 year olds); Asian (vs NH white); with access to health care; in a health setting as worker/volunteer within the previous 12 months; hypertensive, high cholesterol, arthritis or chronic liver disease patient; and, remarkably, unemployed within the 12 months before interview. Moreover, for those who had access to health care within previous 12 months, odds of accepting influenza vaccine increased as the participant was more satisfied with health care within the same period.

Table 2. Odds Ratio Estimates for the Association between Uptake of Seasonal Influenza Vaccine and Use of Internet for Health Purposes among Adults Aged 18-64 years (N~224.1x10[6]), USA, 2014

Exposure	Odds ratio Crude	Odds ratio Adjusted(SE)	95% C.I	P-value
Looked up health information on Internet	1.23	1.11(0.04)	1.02, 1.20	0.01
Filled prescription on Internet	2.48	1.48(0.12)	1.25, 1.75	0.00
Scheduled medical appointment on Internet	1.70	1.18(0.09)	1.02, 1.38	0.03
Communicated with health provider by email	2.05	1.27(0.10)	1.09, 1.48	0.00
Used online group chat to learn about health topics	1.10	0.86(0.09)	0.70, 1.07	0.19
Age (years)				
18-29 (reference)				
30-49	1.41	1.22(0.07)	1.09, 1.37	0.00
50-64	2.28	1.43(0.09)	1.25, 1.64	0.00
65-74	5.58	1.78(0.16)	1.49, 2.12	0.00
75-84	7.35	2.01(0.24)	1.59, 2.53	0.00
85+	9.63	2.50(0.44)	1.76, 3.55	0.00
Sex				
Male (reference)				
Female (Without work/volunteer in a health setting)	1.40	1.29(0.06)	1.17, 1.41	0.00
Female (With work/volunteer in a health setting)	0.67	0.83(0.11)	0.67, 1.08	0.17
Marriage/cohabitation				
Current (reference)				
Previous	1.40	0.86(0.04)	0.79, 0.94	0.00
Never	0.53	0.77(0.06)	0.65, 0.91	0.00

Table 2. Continued

Exposure	Odds ratio Crude	Adjusted(SE)	95% C.I	P-value
Race/Ethnicity				
NH white (reference)				
Black/African American	0.62	0.75(0.04)	0.68, 0.82	0.00
Hispanic	0.53	0.95(0.05)	0.86, 1.04	0.28
Asian	0.98	1.42(0.11)	1.22, 1.64	0.00
Other	0.66	0.93(0.13)	0.71, 1.23	0.62
Region				
Northeast (reference)				
Midwest	0.95	0.98(0.06)	0.87, 1.10	0.76
South	0.87	0.96(0.05)	0.86, 1.06	0.43
West	0.81	0.84(0.05)	0.75, 0.95	0.00
Family income-to-poverty level ratio				
≥2 (≥200% FPL) (reference)				
1-1.99 (100-199% FPL)	0.68	0.78(0.04)	0.71, 0.86	0.00
<1 (<100% FPL)	0.56	0.81(0.04)	0.73, 0.91	0.00
Work status (past 12 months)				
Had job (reference)				
Had no job	1.89	1.29(0.09)	1.12, 1.47	0.00
Worked/volunteered in health setting (past 12 months)				
Male	2.64	5.49(1.00)	3.82, 7.87	0.00
Female	2.15	3.56(0.51)	2.61, 4.71	0.00
Had pneumonia shot	1.34	3.36(0.79)	2.11, 5.35	0.00
Last talked to a health professional				
< 1 year (reference)				
1-2 years	0.29	0.64(0.05)	0.55, 0.75	0.00
2-5 years	0.14	0.37(0.04)	0.30, 0.47	0.00
>5 years/never	0.16	0.43(0.05)	0.34, 0.55	0.00

Table 2. Continued

Exposure	Odds ratio Crude	Adjusted(SE)	95% C.I	P-value
Satisfaction with health care (past 12 months)				
Had no care (reference)				
Very satisfied	7.09	2.15(0.18)	1.81, 2.54	0.00
Somewhat satisfied	4.77	1.65(0.15)	1.37, 1.98	0.00
Not satisfied	3.83	1.38(0.16)	1.11, 1.72	0.00
Smoker				
Never (reference)				
Current	0.57	0.69(0.03)	0.63, 0.76	0.00
Former	1.58	1.08(0.05)	0.10, 1.19	0.10
Arthritis	2.32	1.11(0.05)	1.02, 1.21	0.02
Chronic liver disease	1.84	1.42(0.19)	1.08, 1.87	0.01
High cholesterol	2.29	1.25(0.05)	1.16, 1.35	0.00
Hypertension	2.19	1.11(0.04)	1.03, 1.20	0.01
Had pneumonia shot	5.52	1.93(0.29)	1.43, 2.62	0.00
Without work/volunteer in a health setting	5.52	1.93(0.29)	1.43, 2.62	0.00
With work/volunteer in a health setting	2.24	1.18(0.24)	0.79, 1.77	0.41

The multivariate logistic regression analysis for older adults (Table 3) (F 30, 271=32.38, P < 0.0001) included 8,125 individuals (complete cases) who represented an estimated 94% of the population of those 65 years or older. Higher odds of uptake of influenza vaccine were associated with filling prescriptions on Internet (OR=1.75, C.I: 1.14, 2.69) and communicating with health providers by email (OR=1.48, C.I: 1.01, 2.20). Looking up health information online (OR=1.17, C.I: 0.96, 1.42), scheduling medical appointments over the Internet (OR=3.71, C.I: 0.58, 23.69) and using online group chats to learn about health topics (OR=0.80, C.I: 0.45, 1.41), were not significantly associated with odds of uptake of influenza vaccine among older adults. Age 85 years or older, Asian race/ethnicity, hypertensive, receiving health care within past 12 months (derived from the "satisfaction with health care" variable), and taking a pneumonia shot (interaction with the family income-to-need ratio variable) were associated with increased odds of taking influenza vaccine for older adults. Similar to the analysis for all the adults, odds of taking influenza vaccine were higher the more satisfied the individual was, with healthcare within past 12 months. Associated with decreased odds of accepting seasonal influenza vaccine were black and other races, previous marriage, and family income between 100% and 200% of need. Note that being never married and having family income below the federal poverty level were also negatively associated with odds of using the vaccine among older adults, just like in the general adult population. However, the association was not found to be statistically significant. Also, when the family income-to-need ratio variable modified the effect of the variable for having ever taken pneumonia shot, odds of accepting seasonal influenza vaccine increased significantly for participants who reported ever taking pneumonia shot, for all categories of family income-to-need ratio.

Table 3. Odds Ratio Estimates for the Association between Uptake of Seasonal Influenza Vaccine and Use of Internet for Health Purposes among Adults Aged 65+ years (N~$42.2x10^6$), USA, 2014

Exposure	Odds ratio Crude	Odds ratio Adjusted(SE)	95% C.I	P-value
Looked up health information on Internet	1.42	1.17(0.12)	0.96, 1.42	0.13
Filled prescription on Internet	2.35	1.75(0.38)	1.14, 2.69	0.01
Scheduled medical appointment on the Internet	1.92	3.71(3.44)	0.58, 23.69	0.17
Communicated with health provider by email	2.29	1.48(0.29)	1.01, 2.20	0.049
Used online group chat to learn about health topics	1.03	0.80(0.23)	0.45, 1.41	0.43
Age (years)				
65-74 (reference)				
75-84	1.32	1.13(0.09)	0.97, 1.32	0.12
85+	1.73	1.80(0.23)	1.40, 2.31	0.00
Marriage/cohabitation				
Current (reference)				
Previous	0.85	0.81(0.06)	0.69, 0.94	0.01
Never	0.70	0.65(0.18)	0.37, 1.13	0.13
Race/ethnicity				
NH White (reference)				
Black/African American	0.52	0.63(0.06)	0.51, 0.77	0.00
Hispanic	0.59	0.90(0.11)	0.71, 1.13	0.36
Asian	1.00	1.52(0.28)	1.05, 2.18	0.03
Other	0.57	0.52(0.12)	0.33, 0.81	0.00
Family income-to-poverty ratio				
>2 (>200% FPL) (reference)				
1-2 (100-200% FPL) and never had pneumonia shot	0.68	0.63(0.08)	0.50, 0.81	0.00
1-2 (100-200% FPL) and had a pneumonia shot	0.86	1.01(0.13)	0.78, 1.30	0.96
<1 (<100% FPL) and never had pneumonia shot	0.58	1.00(0.15)	0.74, 1.35	0.99
<1 (<100% FPL) and had a pneumonia shot	0.57	0.74(0.11)	0.54, 1.02	0.07

Table 3. Continued

Exposure	Odds ratio Crude	Adjusted(SE)	95% C.I	P-value
Satisfaction with health care (past 12 months)				
Had no care (reference)				
Very satisfied	10.85	5.68(1.28)	3.64, 8.86	0.00
Somewhat satisfied	8.38	4.84(1.14)	3.04, 7.68	0.00
Not satisfied	6.18	3.57(1.05)	1.99, 6.38	0.00
High cholesterol	1.47	1.17(0.09)	0.99, 1.37	0.05
Hypertension	1.68	1.37(0.11)	1.17, 1.60	0.00
Pneumonia shot ever?				
Family income >200% of FPL	6.58	5.16(0.47)	4.32, 6.16	0.00
Family income 100-200% of FPL	9.13	8.21(1.19)	6.18, 10.91	0.00
Family income < 100% of FPL	4.21	3.84(0.70)	2.68, 5.49	0.00
Response by proxy	0.73	0.57(0.14)	0.36, 0.91	0.02

5 DISCUSSION AND CONCLUSION

The study's main objective was to investigate the relationship between acceptance of seasonal influenza vaccine and use of Internet for health purposes among adults in the United States. Previous literature on the topic is limited both in quantity and the scope of health-related Internet activity used, and American adults are among higher users of both the Internet and influenza vaccine. Analysis of the 2014 NHIS data revealed increased odds of taking the vaccine within 12 months preceding interview among adults who had: searched health information on the Internet, filled prescriptions online, scheduled medical appointments on the Internet, or emailed health provider, all within the same time period. When analysis was limited to older adults who are known to use the Internet less and accept the influenza vaccine more than the general adult population in the US, odds of taking the vaccine also increased for those who had filled prescriptions online and those who communicated with health providers by email. Searching health information on the Internet and scheduling medical appointments online among older adults, as well as use of online group chat to learn about health topics when both adult groups were considered, had no significant association with odds of influenza vaccine uptake. These findings agree, partly, with the study's first hypothesis - each of all forms of health-related Internet use other than use of online group chat to learn about health topics, is associated with increased odds of accepting seasonal influenza vaccine. The exception is in the case of older adults where increased odds of uptake of influenza vaccine was significantly associated only with filling prescriptions online and communicating with health providers via email.

This study is the first of its kind to have examined the same association using very broad scope of health-related Internet activities as they are usually performed by the general public, while looking at their effect on influenza vaccine uptake individually. Among the only two studies to date on the relationship between seasonal influenza vaccine uptake and the use of Internet for health purposes, one (Lau et al., 2012) involved a number of health-related Internet activities similar to the ones used in this study, although using their combined effect, also reported a small but significant increased likelihood of accepting flu vaccine for the Internet users. The other found no significant association between using online personal health record and flu vaccine uptake (Bourgeois et al., 2008). Small sample size, though, could have limited the usability of the findings from this later study as the authors themselves concluded. In any case, a great deal of caution must be exercised when comparing this study and the two mentioned above, in view of the differences in the type of information on influenza and the vaccine to prevent it that the participants could have obtained from the Internet. Information on influenza in the two studies was provided to subjects deliberately, rather than letting them search it themselves. In addition, it was also accurate and generally supportive of use of the vaccine. However, participants in this study likely had come across both positive and negative messages on influenza vaccination.

This study is not without limitations. One of them was a lack of information on some potential confounders that could be important, which were either not included in the survey at all or not released for use by the public. Examples include respondents' education, personal income, frequency of contact with health professionals/providers, and whether activities such as searching health information online were carried out for the benefit of the participants themselves or for someone else. Searching health information for someone else, for example, was common among American adults according to one report (Pew Research Center, 2015). Perhaps at least one of these omitted variables could have explained some of the findings from this study, given the small differences in odds observed in most cases. Another limitation was that the data, being from cross-sectional survey, does not allow determination of order of occurrence of exposure and outcome temporally. As a result, findings were only correlates and causality cannot be claimed. Moreover, the multipurpose nature of the NHIS made information collected on health-related Internet use too broad, with no focus on any specific type of health issue. Without knowing if an individual's Internet use was actually related to influenza vaccine, and whether or not performing that activity had influenced action on vaccine use, there could be high chance for the exposures to be inaccurately defined. Misclassification of both exposure and outcome could have also occurred from using self (or proxy) report to assess them. This method could be prone to problems with accuracy due to poor recall and other similar issues. Furthermore, although about 94% of adults in the NHIS data set were included in the final logistic regression models used to derive inferences for this study, handling missing data by list wise deletion could have resulted in underestimation of standard errors of estimates (Schenker et al. 2006). In other words, the difference in odds observed could have been less. Finally, small sample size could be an issue in the older adult analysis, given the larger standard errors and wider confidence intervals of odds ratio estimates there.

In spite of the limitations, this study has strength in many areas. It involved the largest sample of its kind to date. Also, the sample used is representative of the entire population of noninstitutionalized US adults, as such the findings are generalized to them. The logistic regression model used for odds ratio and standard error estimation adjusted for many confounders. Furthermore, the data used in the analysis is of high quality, considering the stringent measures taken by the producers to ensure its accuracy.

With regards to prevention of seasonal influenza, the study provides another reason to support, at least, some forms of use of the Internet for health purposes as important determinants for vaccine uptake. As the Internet is likely to become more accessible and its use for health purposes more popular because of the potential benefits, policy makers need to pay serious attention to it in any future plans to control influenza. In spite of the study's limitations, a case could be made for emailing health providers and booking appointments with them online as potential interventions. This is because of the observed (and predictably) positive impact on uptake of influenza vaccine attributed to them, even where the association was not statistically significant. The results also support promoting online search for health information by the public, in view of its positive association with odds of influenza vaccination among the general adult group. However, the inconsistency of findings with the older adults and the modest difference in odds even among the general adults, suggests a need to continue to improve the content of the Internet with respect to influenza vaccination. The web should be saturated with information that is not only correct and supportive of influenza vaccine use, but also less technical and easy to access and use by the public. Needless to mention that the special relevance of the older adult sub group in this regard, could hardly be overemphasized. Although filling prescriptions online was found to have significant association with odds of influenza vaccine uptake across all age groups in this study, its practical significance to influenza control, if any, may not be substantial at present. Currently, no form of the vaccine is in use that could be ordered online and shipped to individual consumers, without violating the stringent requirements for shipping and handling necessary for most vaccines (CDC, 2003). In addition, no influenza vaccine has yet been approved for self-administration (Grohskopf et al., 2015). But novel vaccine types and vaccination techniques that are still being researched, e.g. micro needle patches, may change the scenario and make it possible for influenza vaccine to be deliverable to individuals for self-administration (Norman et al., 2014; Kommareddy et al., 2012; Garmise, Staats & Hickey, 2007). Also, future research may be able to establish the mechanism by which filling prescriptions online could have caused the improvement in vaccination rates observed in this study and, possibly, an easier immunization schedule or lower cost for online prescription fillers.

Research efforts should focus on better understanding the relationship between influenza vaccine uptake and health-related Internet use, as well as the usefulness of specific types of public health intervention, such as the ones recommended above. Focus should be on using longitudinal data, better definition, and measurement of exposure and outcome, improved control of confounding and larger sample size for minority groups, such as older adults. Of importance also, is the need to confirm the consistency of some findings from this analysis and/or understand the reason behind them. Examples include the unexpected increase in odds of using influenza vaccine among all adults who reported having no job within the period of interest, and the effect modification role of scheduling medical appointment on Internet on the negative association between odds of accepting the vaccine and searching health information online, among older adults.

REFERENCES

Barker, W. H., & Mullooly, J. P. (1980). Impact of epidemic type A influenza in a defined adult population. *American Journal of Epidemiology*, 112(6), 798-813.

Beckjord, E., Rutten, L. F., Squiers, L., Arora, N., Volckmann, L., Moser, R., & Hesse, B. (2007). Use of the internet to communicate with health care providers in the United States: estimates from the 2003 and 2005 Health Information National Trends Surveys (HINTS). *Journal of Medical Internet Research*, *9*(3), e20

Bourgeois, F. T., Simons, W. W., Olson, K., Brownstein, J. S., & Mandl, K. D. (2008). Evaluation of influenza prevention in the workplace using a personally controlled health record: Randomized controlled trial. *Journal of Medical Internet Research*, 10(1), e5

Centers for Disease Control and Prevention (2003). Guidelines for maintaining and managing the vaccine cold chain. *Morbidity and Mortality Weekly Report*, *52*(42), 1023.

Centers for Disease Control and Prevention (2015a). Influenza (Flu): Types of influenza viruses. http://www.cdc.gov/flu/about/viruses/types.htm. Accessed 7/13/2016

Centers for Disease Control and Prevention (2015b). Influenza vaccination: A summary for clinicians. http://www.cdc.gov/flu/professionals/vaccination/vax-summary.htm

Centers for Disease Control and Prevention (2015c). National early season flu vaccination coverage, United States, November 2014. http://www.cdc.gov/flu/fluvaxview/nifs-estimates-nov2014.htm. Accessed 1/2/2016

Centers for Disease Control and Prevention (2015d). National early season flu vaccination Coverage, United States, November 2015. http://www.cdc.gov/flu/pdf/fluvaxview/nifs-estimates-nov2015.pdf. Accessed 6/13/2016

Choi, N. (2011). Relationship between health service use and health information technology use among older adults: Analysis of the US National Health Interview Survey. *Journal of Medical Internet Research*, 13(2), e33. http://doi.org/10.2196/jmir.1753

Clements, K. M., Chancellor, J., Nichol, K., DeLong, K., & Thompson, D. (2011). Cost-effectiveness of a recommendation of universal mass vaccination for seasonal influenza in the United States. *Value in Health*, 14(6), 800-811.

Cline, R. J., & Haynes, K. M. (2001). Consumer health information seeking on the Internet: The state of the art. *Health Education Research*, 16(6), 671-692.

Cohen, R. A., & Adams, P. F. (2011). Use of the Internet for health information: United States, 2009. *National Center for Health Statistics Data Brief*, (66), 1-7.

Coiera, E., Lau, A. Y., Anvari, F., & Sacks, G. (2010). Healthy.me: An online research platform to support consumer health management in Melbourne. Proceedings of the 18th *Annual Health Informatics Conference*: Informing the business of healthcare, 24-26 August 2010, Melbourne Convention and Exhibition Centre, Brunswick East, Australia. http://search.informit.com.au/documentSummary;dn=429602766943080;res=IELHEA> ISBN: 9780980552027

Covolo, L., Mascaretti, S., Caruana, A., Orizio, G., Caimi, L., & Gelatti, U. (2013). How has the flu virus infected the Web? 2010 influenza and vaccine information available on the Internet. *BMC Public Health*, 13(1), 83.

Fiore, A. E., Uyeki, T. M., Broder, K., Finelli, L., Euler, G. L., Singleton, J. A., ... & Cox, N. J. (2010). Prevention and control of seasonal influenza with vaccines: Recommendations of the Advisory Committee on Immunization Practices (ACIP), 2010. *Morbidity and Mortality Weekly Report*, 59(8), 1–62

Garmise, R. J., Staats, H. F., & Hickey, A. J. (2007). Novel dry powder preparations of whole inactivated influenza virus for nasal vaccination. *AAPS PharmSciTech*, *8*(4), 2-10.

Gregory-Head, B. (1998). Patients and the Internet: Guidance for evidence-based choices. *The Journal of the American College of Dentists*, 66(2), 46-50.

Grohskopf, L. A., Sokolow, L. Z., Olsen, S. J., Bresee, J. S., Broder, K. R., & Karron, R. A. (2015). Prevention and control of influenza with vaccines: Recommendations of the Advisory Committee on Immunization Practices, United States, 2015-16 influenza season. *Morbidity and Mortality Weekly Report*, 64(30), 818-825.

Groves, R. M. (2004). Survey errors and survey costs. Hoboken, NJ: John Wiley & Sons.

He, Y., Zaslavsky, A. M., Landrum, M. B., Harrington, D. P., & Catalano, P. (2009). Multiple imputation in a large-scale complex survey: A practical guide. *Statistical Methods in Medical Research*, 19(6), 653-70.

Heeringa, S. G., West, B. T., & Berglund, P. A. (2010). Applied survey data analysis. Boca Raton, FL: Chapman & Hall/CRC Press.

Hosmer, D. W., Lemeshow, S., & Cook, E. D. (2000). Applied logistic regression: Second edition. Hoboken, NJ: John Wiley & Sons.

Kata, A. (2010). A postmodern Pandora's Box: Anti-vaccination misinformation on the Internet. *Vaccine*, 28(7), 1709-1716.

Kommareddy, S., Baudner, B. C., Oh, S., Kwon, S. Y., Singh, M., & O'hagan, D. T. (2012). Dissolvable microneedle patches for the delivery of cell-culture-derived influenza vaccine antigens. *Journal of Pharmaceutical Sciences*, *101*(3), 1021-1027.

LaPerriere, B., Edwards, P., Romeder, J. M., & Maxwell-Young, L. (1998). Using the Internet to support self-care. *The Canadian Nurse*, 94(5), 47-48.

Lau, A. Y. S., Sintchenko, V., Crimmins, J., Magrabi, F., Gallego, B., & Coiera, E. (2012). Impact of a Web-based personally controlled health management system on influenza vaccination and health services utilization rates: A randomized controlled trial. *Journal of the American Medical Informatics Association*, 19(5), 719–727

Lau, A. Y., Arguel, A., Dennis, S., Liaw, S. T., & Coiera, E. (2015). “Why didn’t it work?” Lessons from a randomized controlled trial of a Web-based Personally Controlled Health Management System for Adults with Asthma. *Journal of Medical Internet Research*, 17(12), e283

Lenhart, A., Purcell, K., & Smith, A. (2010). Social media and mobile Internet use among teens and young adults. Pew Internet and American Life Project. http://www.pewinternet.org/files/oldmedia/Files/Reports/2010/PIP_Social_Media_and_Young_Adults_Report_Final_with_toplines.pdf. Accessed 07/25/2016

Levy, P. S., & Lemeshow, S. (2013). Sampling of populations: Methods and applications. Hoboken, NJ: John Wiley & Sons.

Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, 9(1), 1-19.

Martinello, R. A., Jones, L., & Topal, J. E. (2003). Correlation between healthcare workers' knowledge of influenza vaccine and vaccine receipt. *Infection Control*, 24(11), 845-847.

Molinari, N. A. M., Ortega-Sanchez, I. R., Messonnier, M. L., Thompson, W. W., Wortley, P. M., Weintraub, E., & Bridges, C. B. (2007). The annual impact of seasonal influenza in the US: Measuring disease burden and costs. *Vaccine*, 25(27), 5086-5096.

Monto A.S., Gravenstein S. and Elliot M. (2000). Clinical signs and symptoms predicting influenza infection. *Archives of Internal Medicine*, 160, 3243-3247

Moriarity, C., & Dahlhamer, J. (2012). Adjustment for unit nonresponse in the National Health Interview Survey. Proceedings of the Joint Statistical Meetings. Alexandria, VA: *American Statistical Association*, 4783-4789.

Mortimer, N. J., Rhee, J., Guy, R., Hayen, A., & Lau, A. Y. (2015). A Web-based Personally Controlled Health Management System increases sexually transmitted infection screening rates in young people: A randomized controlled trial. *Journal of the American Medical Informatics Association*, 22, 805-814

National Center for Health Statistics (2015a). Multiple imputation of family income and personal earnings in the National Health Interview Survey: Methods and examples. https://www.cdc.gov/nchs/data/nhis/tecdoc14.pdf. Accessed 07/25/2016

National Center for Health Statistics (2015b). National Health Interview Survey, 2014. Public-use data file and documentation. Survey description. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/NHIS/2014/srvydesc.pdf

Nelson K.E. & Williams C.M. (2014). Infectious disease epidemiology: Theory and practice. 3rd edition. Burlington, MA: Johns & Barnett.

Nichol, K. L., Lofgren, R. P., & Gapinski, J. (1992). Influenza vaccination: knowledge, attitudes, and behavior among high-risk outpatients. *Archives of Internal Medicine*, 152(1), 106-110

Nichol, K. L., Mac Donald, R., & Hauge, M. (1996). Factors associated with influenza and pneumococcal vaccination behavior among high-risk adults. *Journal of General Internal Medicine,* 11(11), 673-677.

Norman, J. J., Arya, J. M., McClain, M. A., Frew, P. M., Meltzer, M. I., & Prausnitz, M. R. (2014). Microneedle patches: usability and acceptability for self-vaccination against influenza. *Vaccine*, *32*(16), 1856-1862.

Parsons V. & Moriarity C. (2007). Review of NHIS public-design structures. Proceedings of the Joint Statistical Meetings. Alexandria, VA: *American Statistical Association*, 2903–2909

Parsons, V. L., Moriarity, C., Jonas, K., Moore, T. F., Davis, K. E., & Tompkins, L. (2014). Design and estimation for the national health interview survey, 2006-2015. *Vital and Health Statistics,* 2(165), 1-53.

Pew Research Center (2000). Internet and American life project: African-Americans and the Internet. http://www.pewinternet.org. Accessed 12/25/2015

Pew Research Center (2015). Health facts. http://www.pewinternet.org/fact-sheets/health-fact-sheet. Accessed on 12/25/2015

PricewaterhouseCoopers (1999). Net to radically change health care industry. http://www.nua.net.surveys. Accessed on 10/10/2001

Poland, G. A., & Jacobson, R. M. (2001). Understanding those who do not understand: a brief review of the anti-vaccine movement. *Vaccine*, 19(17), 2440-2445.

Rao, J. N. K., & Scott, A. J. (1987). On simple adjustments to chi-square tests with sample survey data. *The Annals of Statistics*, 15(1), 385-397

Risk, A., & Dzenowagis, J. (2001). Review of Internet health information quality initiatives. *Journal of Medical Internet Research*, 3(4), e28

Robinson, T. N., Patrick, K., Eng, T. R., & Gustafson, D. (1998). An evidence-based approach to interactive health communication: a challenge to medicine in the information age. *Journal of the American Medical Association*, 280(14), 1264-1269.

Schenker, N., Raghunathan, T. E., Chiu, P. L., Makuc, D. M., Zhang, G., & Cohen, A. J. (2006). Multiple imputation of missing income data in the National Health Interview Survey. *Journal of the American Statistical Association*, 101(475), 924-933.

Sonnenberg, F. A. (1997). Health information on the Internet: Opportunities and pitfalls. *Archives of Internal Medicine*, 157(2), 151-152.

Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., … & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *British Medical Journal*, *338*, b2393.

Texas A&M University Division of Research (2012). Guidance on public use data sets. https://vpr.tamu.edu/compliance/rcc/irb/irb-guidance/public-use-data-sets. Accessed on 7/21/2016

Thompson, M. G., Shay, D. K., Zhou, H., Bridges, C. B., Cheng, P. Y., Burns, E., ... & Cox, N. J. (2010). Estimates of deaths associated with seasonal influenza-United States, 1976-2007. *Morbidity and Mortality Weekly Report*, 59(33), 1057-1062.

United Nations Statistical Division (2008). Designing household survey samples: practical guidelines. Volume F/98. New York, NY: United Nations.

Weng, H. Y., Hsueh, Y. H., Messam, L. L. M., & Hertz-Picciotto, I. (2009). Methods of covariate selection: directed acyclic graphs and the change-in-estimate procedure. *American Journal of Epidemiology*, 169(10), 1182-1190.

Wolfe, R. M., Sharp, L. K., & Lipsky, M. S. (2002). Content and design attributes of anti-vaccination web sites. *Journal of The American Medical Association*, 287(24), 3245-3248

Wolfe, R. M., & Sharp, L. K. (2005). Vaccination or immunization? The impact of search terms on the Internet. *Journal of Health Communication,* 10(6), 537-551.